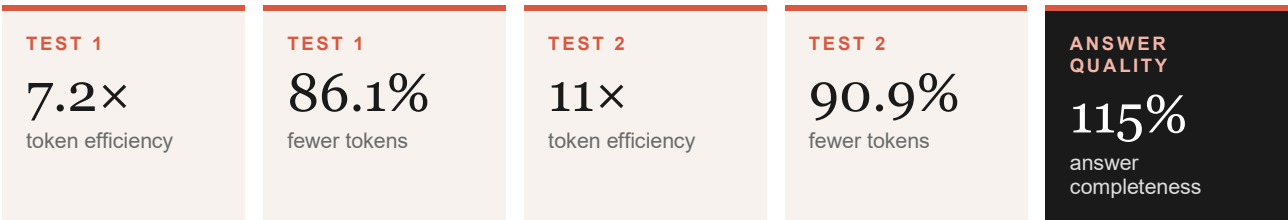


Same AI. Same data.

7.2× to 11× more efficient.

Traditional MCP workflows versus the Holiday Environment

A controlled comparison of how AI retrieves information from connected business systems. The model and the underlying records stayed the same; only the access architecture changed.



KEY FINDINGS

- Holiday answered the same customer question with **86.1% to 90.9% fewer tokens** than a traditional MCP workflow.
- It needed **2 agent-facing calls** in both tests. The traditional workflow needed 5, then 8 as the task expanded.
- Holiday returned **every required data point plus linked context**, scoring 115% on answer completeness.
- The gap grew as the workflow expanded, which suggests larger deployments will see larger savings.

01 / THE TEST

The business question

A customer sends an email. The AI must connect that correspondence to the correct business in the CRM, review relevant history, inspect active jobs and retrieve the related invoices from the accounting system. A complete answer reconciles information across systems and highlights where records disagree.

Four connected business systems

System	What the AI needed to establish
Inbox / email	Identify the sender, read relevant correspondence and recover commitments or changes.
CRM	Identify the contact and company, even where the email address differs from accounting records.
Jobs / project management	Identify work associated with that customer and establish project status.
Accounting	Identify related invoices, amounts and payment position.

Test environment

The benchmark ran on Holiday's own business rather than a synthetic dataset:

- **Live systems with real customer data:** Outlook (inbox), Xero (accounting) and Holiday's CRM.
- **Test environment:** job and project records, standing in for a project management system.
- **Same model throughout:** GPT-5.6 Sol, with identical underlying records for both workflows.
- **Only the architecture changed:** the traditional workflow navigated each system through separate MCP tools; Holiday answered from its connected business environment.

Customer details from live systems are not reproduced in this report.

What we measured

Measure	What it tells you
Agent-facing API calls	How many separate tool calls the AI had to make to assemble the answer.
Total tokens processed	The volume of text the model had to read and write. Drives AI cost and response time.
Answer completeness	How much of the required information was returned, plus any linked context. See section 04.

02 / TEST 1: FIRST CUSTOMER-CONTEXT TASK

7.2× more efficient on the same question

Measure	Traditional MCP workflow	Holiday Environment
Agent-facing API calls	5	2
Total tokens processed	268,625	37,295
Reduction relative to MCP	—	86.1%
Token efficiency	Baseline	7.2× more efficient

TOTAL TOKENS PROCESSED, TEST 1

Traditional MCP		268,625
Holiday Environment		37,295

Both responses identified the customer, the associated projects and invoices, and surfaced mismatches between correspondence and job records. In qualitative review, Holiday delivered the clearer overall business narrative; the traditional workflow supplied some more granular accounting fields. Answer completeness was scored separately, using the method in section 04.

What happens as the workflow expands?

The second experiment retrieved the same customer information from the same four systems, but the traditional MCP workflow needed eight agent-facing calls. Holiday completed the identical information task with the same two business-context calls.

Measure	Traditional MCP workflow	Holiday Environment
Systems involved	Inbox, CRM, accounting, jobs	Inbox, CRM, accounting, jobs
Agent-facing API calls	8	2
Input tokens	409,937	36,652
Cached input tokens	354,048	0
Output tokens	1,040	643
Total tokens	410,977	37,295
Tokens saved	—	373,682
Token efficiency	Baseline	11.0× more efficient (90.9% fewer tokens)

TOTAL TOKENS PROCESSED, TEST 2

Traditional MCP		410,977
Holiday Environment		37,295

Traditional MCP calls recorded in test 2

System	Calls	Operations
Inbox	1	emails.list
CRM	2	clients.list, client
Accounting	2	xero-contacts.list, accounting-documents.list
Jobs	3	jobs.clients.list, jobs.client, jobs.list

Because the expanded run sought the same customer information, the same two-call Holiday result (37,295 tokens) is the comparison baseline. The traditional workflow consumed 373,682 additional tokens, even though 354,048 of its input tokens were cached.

A note on caching. Cached tokens are a subset of input tokens. Cached input tokens remain part of total token usage but are billed at reduced rates, depending on the model provider.

How we score answer completeness

Token efficiency only matters if the answer is still complete. To measure that, we count data points: the individual facts or records an answer contains.

THE FORMULA

Completeness = relevant data points returned ÷ data points required

The required data points are everything a complete, correct answer must contain. Relevant data points include those, plus linked records that add useful context.

Reading the score

Score	Meaning	Example: “find me the invoices” (7 exist)
Below 100%	Required records were missed.	6 of 7 invoices returned: 86%
100%	Every required record was returned.	All 7 invoices returned: 100%
Above 100%	Everything required, plus relevant linked context.	7 invoices + the 2 emails that sent them: 130%

Worked example: quality and efficiency together

Completeness and token use can be combined into a single efficiency figure: useful data points per token. In this illustrative example, four workflows answer the same question, where 7 invoices exist.

Workflow	Tokens	Returned	Completeness	Data points per token
A	10	7 invoices	100%	0.70
B	5	7 invoices + 7 sent emails	200%	2.80
C (smaller model)	8	6 invoices	86%	0.75
D	10	7 invoices + 7 sent emails + 7 related jobs	300%	2.10

Workflow B is the most efficient: it returns a more useful answer than A while using half the tokens. Workflow C is cheaper than A but incomplete. Token counts in this example are illustrative.

Benchmark result

	Data points required	Relevant data points returned	Completeness
Holiday Environment	67	77 (all 67 required + 10 linked)	114.9%
Traditional MCP workflow	67	67	100%

Holiday returned every data point a complete answer required, plus ten linked records such as the correspondence and jobs connected to the customer's invoices. Combined with the token results, Holiday delivered a more complete answer at a fraction of the processing cost.

Why the Environment matters

The Holiday Environment provides a unified discovery layer and connected customer context. Instead of asking the model to navigate separate tools and repeatedly rebuild the relationship between an inbox sender, a CRM record, a job and an invoice, the Environment exposes those business relationships through a common, permission-controlled interface. Cross-system answers become easier to assemble and easier to audit.

Governance is built around controlled data access and traceable activity: business reads are permission-filtered, and Holiday maintains records of agent and business activity. The value is not only efficiency. Organisations also gain one consistent place to supervise which information an AI can use and how it reached an answer.

Why this is a conservative starting point

The tested traditional-style catalogue exposed only about 60 APIs. Production deployments spanning systems such as Outlook, Salesforce, Odoo and payroll may offer hundreds of operations collectively. A larger catalogue introduces more tool-selection and identity-matching work. The exact token impact depends on how that catalogue is presented to the model, but this test does not represent the upper limit of multi-system complexity.

Conclusion

Two experiments using GPT-5.6 Sol and the same underlying data showed that the Holiday Environment returned actionable customer knowledge with 7.2× and 11× greater token efficiency than traditional MCP workflows, while scoring 115% on answer completeness.

The expanded comparison required 75% fewer agent-facing calls and 90.9% fewer tokens. Because traditional agents must navigate independent providers, tool catalogues and identity links, we expect Holiday's unified business relationships to deliver even greater value as more systems are added to a workflow.